AD-A078 960 PRINCETON UNIV NJ DEPT OF STATISTICS F/G 12/1

ASYMPTOTIC INFERENCE FOR EIGENVECTORS.(U)

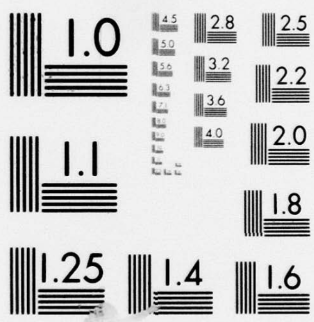
OCT 79 D E TYLER N00014-67-0151-0017

UNCLASSIFIED TR-157-SER-2 NL

| OF |

AD
A078960

END
DATE
FILMED
1-80
DDC




MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A

6

# ASYMPTOTIC INFERENCE FOR EIGENVECTORS

David E. Tyler

Technical Report #157, Series 2

Department of Statistics
Princeton University

October, 1979

27





Research supported in part by ONR Grants N0014-67-0151-0017,

and N00014-75-C-0453, ARO Grant DAHC-04-74-G0178 and ERDA Grant E-111-2310.



406873

ABSTRACT

Asymptotic procedures are given for testing certain hypotheses concerning eigenvectors and for constructing confidence regions for eigenvectors. These asymptotic procedures are derived under fairly general conditions on the estimates of the matrix whose eigenvectors are of interest. Applications of the general results to principal components analysis and canonical variate analysis are given.

## 1. INTRODUCTION AND SUMMARY

Let M be a $(p \times p)$ matrix which is symmetric in the metric of the positive definite symmetric matrix $\Gamma$, i.e. $\Gamma M$ is symmetric. Let the eigenvalues of M be represented by $\lambda_1 \geq \lambda_2 \ldots \geq \lambda_p$. Also, let $M_n$ be a sequence of estimates of M such that $a_n(M_n - M)$ converges in distribution to a multivariate normal distribution, where $a_n$ is an increasing sequence of real numbers, and let A be an $(r \times p)$ matrix with $\text{rank}(A) = r$.

In this paper, under the assumption that $\lambda_{i-1} \neq \lambda_i$ and $\lambda_{i+m-1} \neq \lambda_{i+m}$, the following null hypothesis is considered.

(1.1) $H_o$: the columns of A lie in the subspace generated by the set of eigenvectors of M associated with the roots $\lambda_i, \lambda_{i+1}, \ldots, \lambda_{i+m-1}$.

The assumption on the eigenvalues is to be interpreted as $\lambda_{i+m-1} \neq \lambda_{i+m}$ when $i = 1$, and $\lambda_{i-1} \neq \lambda_i$ when $i+m-1 = p$.

Under fairly general condition on $M_n$, a consistent asymptotic chi-square test of $H_o$ is given. This test is based upon the asymptotic normality of the "orthogonal" projection of the columns of A onto the subspace generated by the eigenvectors of $M_n$ associated with the ith to (i+m-1)th roots of $M_n$.

An asymptotic confidence region for the subspace generated by the eigenvectors of M assoicated with the roots $\lambda_i, \lambda_{i+1}, \ldots, \lambda_{i+m-1}$ is then given. This confidence region is based upon the asymptotic chi-square test of $H_o$ for the special case when $r = m$. Furthermore, an asymptotic chi-square test is given for the hypothesis that the subspace generated by the columns of K, a $(k \times p)$ matrix with $\text{rank}(K) = k \geq m$, contains the subspace generated by the eigenvectors of M associated with the roots

$\lambda_i, \lambda_{i+1} \ldots, \lambda_{i+m-1}$ . This test is constructed by relating this hypothesis to a hypothesis of the form given in (1.1).

Anderson (1963) gives an asymptotic chi-square test of $H_o$ for the special case of $m = 1$ and when $M_n$ is the sample covariance matrix from a multivariate normal sample with population covariance matrix M. This paper is thus a generalization of Anderson's results.

James (1977) gives an exact test for a hypothesis similar to (1.1) when $M_n$ is the sample covariance matrix from a multivariate normal sample with population covariance matrix M. James considers the hypothesis that the columns of A generate an invariant subspace of M. His hypothesis does not state with which eigenvalues of M the invariant space is associated. The approach used by James uses special properties of the sample covariance matrix from a normal sample and does not readily generalize to other matrices.

For other related works on the distributional and inferential theory for eigenvectors, the reader is refered to Anderson (1951), Mallows (1961), Chambers (1967), Izenman (1976) and Suguira (1976).

Applications of the general results in this paper are illustrated through the following two examples: the principal component vectors for the covariance matrix of a multivariate normal distribution, and the canonical vectors associated with two random vectors which jointly have a multivariate normal distribution.

## 2. PRELIMINARIES

In this section, let S be a $(q \times q)$ real matrix which is symmetric in the metric of a real positive definite symmetric matrix T. In order to establish notation and vocabulary, the eigenvalue problem for S is briefly reviewed below. A more detailed review can be found in Kato (1966) or Nerring (1970).

If $S\underset{\sim}{x} = \lambda\underset{\sim}{x}$ for some $\underset{\sim}{x} \neq 0$, then $\lambda$ is an eigenvalue of S and $\underset{\sim}{x}$ is an eigenvector of S associated with $\lambda$. All eigenvalues of S are real. The spectral set of S, denoted $\mathcal{S}$, is the set of all eigenvalues of S. Eigenvalues of "symmetric" matrices have the following important continuity property.

LEMMA 2.1. If the $(q \times q)$ matrix $S_k$ is symmetric in the metric of $T_k$, with eigenvalues $\lambda_1(S_k) \geq \lambda_2(S_k) \geq \ldots \geq \lambda_q(S_k)$, and $S_k \to S$ as $k \to \infty$, then $\lambda_j(S_k) \to \lambda_j(S)$ as $k \to \infty$.

The eigenspace of S associated with $\lambda$ is $V(\lambda) = \{\underset{\sim}{x} \in R^q \mid S\underset{\sim}{x} = \lambda\underset{\sim}{x}\}$, where $R^q$ is the set of all q-dimensional real vectors. The dimension of $V(\lambda)$ is the multiplicity of $\lambda$, say $m(\lambda)$. If $\lambda$ and $\mu$ are two distinct eigenvalues of M, then $V(\lambda)$ and $V(\mu)$ are orthogonal subspaces in the metric of T. That is, if $\underset{\sim}{x} \in V(\lambda)$ and $\underset{\sim}{y} \in V(\mu)$, then $\underset{\sim}{x}'T\underset{\sim}{y} = 0$.

Since S is symmetric in the metric of T, we have the decomposition, $R^q = \Sigma_{\lambda \in \mathcal{S}} V(\lambda)$. The eigenprojection of S associated with $\lambda$, denoted $P(\lambda)$, is the projection operator onto $V(\lambda)$ with respect to this decomposition of $R^q$. If $\mathcal{V}$ is any subset of the spectral set $\mathcal{S}$, then the total eigenprojection for S associated with the eigenvalues in $\mathcal{V}$ is defined to be $\Sigma_{\lambda \in \mathcal{V}} P(\lambda)$. For any set of vectors $\{\underset{\sim}{x}_j\}$ in $V(\lambda)$ such that

$\underset{\sim}{x}_j' T \underset{\sim}{x}_k = \delta_{jk}$, where $\delta_{jk}$ denotes the Kronecker delta, $P(\lambda)$ has the representation $P(\lambda) = \Sigma_{j=1}^{m(\lambda)} \underset{\sim}{x}_j \underset{\sim}{x}_j' T$. Thus $P(\lambda)$ is symmetric in the metric of T.

The spectral decomposition of S is $S = \Sigma_{\lambda \epsilon S} \lambda P(\lambda)$. If all the eigenvalues of S are non-negative, then the square root of S is to be defined as $S^{\frac{1}{2}} = \Sigma_{\lambda \epsilon S} \lambda^{\frac{1}{2}} P(\lambda)$.

A generalized inverse of S is any $S^-$ such that $SS^-S = S$. The Moore-Penrcse generalized inverse of S, denoted by $S^+$, can be represented by $S^+ = \Sigma_{\lambda \epsilon S, \lambda \neq 0} \lambda^{-1} P(\lambda)$.

In working with random matrices, it is necessary to introduce the following notation. If B is a $(b \times t)$ matrix, then vec(B) is the transformation of B into a bt-dimensional vector in the following fashion. Let $B = [\underset{\sim}{b}_1 \ \underset{\sim}{b}_2 \ \ldots \ \underset{\sim}{b}_t]$, where $\underset{\sim}{b}_j$ is the jth column of B, then

$$\text{vec}(B) = \begin{bmatrix} \underset{\sim}{b}_1 \\ \underset{\sim}{b}_2 \\ \vdots \\ \underset{\sim}{b}_t \end{bmatrix}. \tag{2.1}$$

If B is a $(b \times t)$ matrix and C is a $(c \times u)$ matrix, then the Kronecker product of B and C is the $(bc \times tu)$ partitioned matrix $B \otimes C = [b_{jk}C]$, $j = 1, 2, \ldots, b$ and $k = 1, 2, \ldots, t$ with j varying over rows of matrices and k varying over columns of matrices.

An important property relating the "vec" transformation and the Kronecker product is

$$\text{vec}(BCD) = (D' \otimes B)\text{vec}(C), \tag{2.2}$$

where the dimensions of the matrices B, C, and D are such that the multiplications are properly defined. Other properties of the "vec" transformation and the Kronecker product can be found in Neudecker (1968, 1969).

The commutation matrix or permuted identity matrix is the $(ab \times ab)$ matrix $I_{(a,b)} = \Sigma_{i=1}^{a}\Sigma_{j=1}^{b}E_{ij} \otimes E_{ij}'$, where $E_{ij}$ is an $(a \times b)$ matrix with a one in the $(i,j)$ position and zeroes elsewhere. The commutation matrix has been extensively investigated recently by Magnus and Neudecker (1979). Two important properties of the commutation matrix are

$$I_{(a,b)}\,\mathrm{vec}(B) = \mathrm{vec}(B'),\ \text{and} \tag{2.3}$$

$$I_{(a,b)}\,(C \otimes D) = (D \otimes C)I_{(c,\,d)}, \tag{2.4}$$

where B is $(b \times a)$, C is $(b \times d)$, and D is $(a \times c)$.

If $\underset{\sim}{Y}$ is a random vector, let $\mathrm{var}(\underset{\sim}{Y})$ represent the covarinace matrix of $\underset{\sim}{Y}$. If B is a random matrix, then for convenience, var[vec(B)] is to be written as var(B).

## 3. ASSUMPTIONS

In order to form an asymptotic test for (1.1), a sequence of estimators $M_n$ for M are needed which satisfies the following assumptions.

ASSUMPTION 3.1.

(i) $M_n$ is symmetric in the metric of $\Gamma_n$, a positive definite symmetric matrix, with $\Gamma_n \to \Gamma$ in probability.

(ii) $a_n(M_n - M) \to N$ in distribution, where $a_n$ is an increasing sequence of positive numbers such that $a_n \to \infty$ as $n \to \infty$, and N is a multivariate normal matrix with zero mean and $\mathrm{var}(N) = \Sigma$.

(iii) For B which is $(p \times p)$, $\Sigma\mathrm{vec}(\Gamma B) = 0$ implies $M(B + B') = 0$.

It is also necessary to have a sequence of estimators $\Sigma_n$ for $\Sigma$ which satisfies the following properties.

ASSUMPTION 3.2

(i) $\Sigma_n$ is symmetric and positive semi-definite.

(ii) $\Sigma_n \to \Sigma$ in probability.

(iii) Let $\Omega_n = \{\Sigma_n \mathrm{vec}(\Gamma_n B) = 0$ implies $M_n(B + B') = 0\}$ then

$\mathrm{Prob}(\Omega_n) \to 1$.

Finally, it is to be understood that the asymptotic procedures given in this paper are only defined on the set

$$C_n = \{\hat{\lambda}_{i-1} \neq \hat{\lambda}_i \text{ and } \hat{\lambda}_{i+m-1} \neq \hat{\lambda}_{i+m}\} , \tag{3.1}$$

where $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \ldots \geq \hat{\lambda}_p$ are the eigenvalues of $M_n$. It is irrelevant to the asymptotic properties of the procedures what action is taken on the complement of $C_n$, since by the continuity of the eigenvalues of "symmetric" matrices, that is Lemma 2.1, $\mathrm{Prob}[C_n] \to 1$.

## 4. ASYMPTOTIC DISTRIBUTION OF THE EIGENPROJECTION

Let $w = \{\lambda_i, \lambda_{i+1}, \ldots, \lambda_{i+m-1}\}$ and let $\hat{w} = \{\hat{\lambda}_{i+1}, \ldots, \hat{\lambda}_{i+m-1}\}$. Also, for $\lambda$ an eigenvalue of M, let $P_\lambda$ represent the eigenprojection of M associated with $\lambda$, and let $Q_\lambda = (M - \lambda I)^+$. For $\lambda$ an eigenvalue of $M_n$, let $\hat{P}_\lambda$ represent the eigenprojection of $M_n$ associated with $\lambda$, and let $\hat{Q}_\lambda = (M_n - \lambda I)^+$. For convenience, define $P_o = \Sigma_{\lambda \epsilon w} P_\lambda$ and $\hat{P}_o = \Sigma_{\lambda \epsilon \hat{w}} \hat{P}_\lambda$. $P_o$ represents the total eigenprojection of M associated with the eigenvalues of $M$ in $w$, and $\hat{P}_o$ represents the total eigenprojection of $M_n$ associated with the eigenvalues of $M_n$ in $\hat{w}$.

The null hypotheses (1.1) can thus be rephrased as

$$H_o : P_o A = A \tag{4.1}$$

where $A$ is $(p \times r)$ with $\mathrm{rank}(A) = r$.

A natural statistic to consider in testing $H_o$ is $(\hat{P}_o A - A)$. In obtaining the asymptotic distribution of this statistic, the Taylor series expansion of $\hat{P}_o$ about $P_o$ is to be used. This expansion is given in the following lemma. The lemma is a simplified version of more general results given in Chapter 2 of Kato (1966). A proof of this simplified version can be found in Appendix B of the author's dissertation.

LEMMA 4.1. Let $d_o = \min\{\lambda_{i-1} - \lambda_i, \lambda_{i+m-1} - \lambda_{i+m}\}$, and $d_1 = (\lambda_i - \lambda_{i+m-1})$. Also define the norm $||B|| = [\max \text{ eigenvalue}(\Gamma^{-1}B'\Gamma B)]^{\frac{1}{2}}$.

If $||M_n - M|| \le d_o/2$, then

$$\hat{P}_o = P_o - \Sigma_{\lambda \epsilon w}[P_\lambda(M_n - M)Q_\lambda + Q_\lambda(M_n - M)P_\lambda] + E_n ,$$

where $||E_n|| \le (1 + d_1/d_o)(2||M_n - M||/d_o)^2(1 - 2||M_n - M||/d_o)^{-1}$.

This lemma immediately yields the following limiting distribution,

$$a_n(\hat{P}_o - P_o) \to N_o = -\Sigma_{\lambda \epsilon w}[P_\lambda N Q_\lambda + Q_\lambda N P_\lambda] \tag{4.2}$$

in distribution. So, under $H_o$, $a_n(\hat{P}_o A - A) \to N_o A$ in distribution, with the covariance matrix of $N_o A$ being

(4.4) $$\mathcal{I}_o(A) = A_o'\mathcal{I}A_o, \text{ where}$$

$$A_o = -[\Sigma_{\lambda\in w}P_\lambda \otimes (I-P_o')Q_\lambda'](A \otimes I) = \Sigma_{\lambda\in w}\Sigma_{\mu\notin w}(\lambda-\mu)^{-1}P_\lambda A \otimes P_\mu' .$$

In the next section, the asymptotic normality of $a_n(\hat{P}_o A - A)$ is to be used in forming an asymptotic chi-square test. Before doing so, the rank of $\mathcal{I}_o(A)$ is needed.

THEOREM 4.2. If the columns of A are in the range of $P_o$, then $\text{rank}[\mathcal{I}_o(A)] = (p-m)r$.

PROOF. The proof consists of determining the null space of $\mathcal{I}_o(A)$. For G which is $(p \times r)$, $[\text{vec}(\Gamma G)]'\mathcal{I}_o(A)\ \text{vec}(\Gamma G) = [\text{vec}(\Gamma B)]'\mathcal{I}\text{vec}(\Gamma B)$, where $B = \Sigma_{\lambda\in w}\Sigma_{\mu\notin w}(\lambda-\mu)^{-1}P_\mu GA'P_\lambda$. So by Assumption 2.1.iii, $\mathcal{I}_o(A)\text{vec}(\Gamma G) = 0$ implies $M(B + B') = 0$, which implies $P_\mu G = 0$ for $\mu \notin w$.

The last implication is justified by the following contrapositive argument. Suppose $\mu \notin w$ and $P_\mu G \neq 0$. Case I: $0 \notin w$. $M(B + B')P_\mu = (\Sigma_{\lambda\in w}\lambda(\lambda-\mu)^{-1}P_\lambda)AG'P_\mu \neq 0$ since $\text{rank}[(\Sigma_{\lambda\in w}\lambda(\lambda-\mu)^{-1}P_\lambda)A = r$. Case II: $0 \in w$. $P_\mu M(B + B') = \mu P_\mu GA'(\Sigma_{\lambda\in w}(\lambda-\mu)^{-1}P_\lambda) \neq 0$

The converse, that is $P_\mu G = 0$ for all $\mu \notin w$ implies $\mathcal{I}_o(A)\text{vec}(\Gamma G) = 0$. is obviously true. Thus, the null space of $\mathcal{I}_o(A)$ is $\eta = \{\text{vec}(\Gamma G) | (I-P_o)G = 0\}$. Thus, $\text{dimension}(\eta) = mr$, and so $\text{rank}[\mathcal{I}_o(A)] = (p-m)r$.

## 5. AN ASYMPTOTIC CHI-SQUARE TEST

In this section, an asymptotic chi-square test of $H_o$ is given which is based upon the asymptotic normality of $a_n(\hat{P}_o A - A)$. Due to the singularity of $\mathcal{I}_o(A)$, the chi-square test and its properties are not straightforward. So, for clarity, most proofs for this section are given in the appendix.

THEOREM 5.1. Let

$$\hat{\Sigma}_o(A) = (A' \otimes I)[\Sigma_{\lambda\epsilon\hat{w}} \hat{P}'_\lambda \otimes \hat{Q}_\lambda(I-\hat{P}_o)]\Sigma_n[\Sigma_{\lambda\epsilon\hat{w}} \hat{P}_\lambda \otimes (I-\hat{P}'_o)\hat{Q}'_\lambda](A \otimes I),$$

and let $[\hat{\Sigma}_o(A)]^-$ represent any generalized inverses of $\hat{\Sigma}_o(A)$. Also, define $T_n(A) = a_n^2\{vec[(\hat{P}_o - I)A]\}'[\hat{\Sigma}_o(A)]^- vec[(\hat{P}_o - I)A]$. Then, under $H_o$, $T_n(A) \to \chi^2_{r(p-m)}$ in distribution.

PROOF. See the appendix.

Theorem 5 states that the limiting distribution of $T_n(A)$ under $H_o$ does not depend upon the choice of generalized inverses for $\hat{\Sigma}_o(A)$. The next theorem states that asymptotically the value of $T_n(A)$ does not depend upon the choice of the generalized inverses for $\hat{\Sigma}_o(A)$.

THEOREM 5.2. Let $T_n(A)$ be defined as in Theorem 5.1.

(i) On the set $\{rank(\hat{P}_o A) = r\}$, $T_n(A)$ is invariant under different choices of a generalized inverse for $\hat{\Sigma}_o(A)$.

(ii) Whether or not $H_o$ is true, $Prob[rank(\hat{P}_o A) = r] \to 1$.

PROOF. See the appendix.

So, on the set $\{rank(\hat{P}_o A) = r\}$, $T_n(A)$ is unique and the Moore-Penrose inverse for $\hat{\Sigma}_o(A)$ can thus be used on this set. In addition, if $M_n$ is symmetric, then $T_n(A)$ has the representation

(5.1) $$T_n(A) = a_n^2[vec(A)]'[\hat{\Sigma}_o(A)]^+ vec(A)$$

on the set $\{\text{rank}(\hat{P}_o A) = r\}$. This statement is justified by noting that for a symmetric matrix B, $B\underset{\sim}{x} = \underset{\sim}{0}$ if and only if $B^+\underset{\sim}{x} = \underset{\sim}{0}$, and it is easy to verify that $\hat{\Sigma}_o(A)\text{vec}(\hat{P}_o A) = \underset{\sim}{0}$.

Now, application of Theorems 5.1 and 5.2 gives the following asymptotic $\alpha$ level test of $H_o$.

(5.2) Reject $H_o$ if

(i) $\text{rank}(\hat{P}_o A) < r$, or

(ii) $\text{rank}(\hat{P}_o A) = r$ and $T_n(A) > \chi^2_{r(p-m)}(\alpha)$,

where $\chi^2_k(\alpha)$ is the $(1-\alpha)$ percentile of a $\chi^2_k$ distribution.

By Theorem 5.2.ii, it is irrelevant to the asymptotic properties of a test of $H_o$ what action is taken on the set $\{\text{rank}(\hat{P}_o A) < r\}$. However, rejecting $H_o$ for this case enables the rejection region to be "continuous" in the sense that for any sequence $\{A_k\}$ such that $\text{rank}(\hat{P}_o A_k) = r$ and $A_k \to A$, where $\text{rank}(\hat{P}_o A) < r$, then

$$T_n(A_k) \to \infty, \text{ as } k \to \infty. \tag{5.3}$$

The proof of property (5.3) is given in the appendix. Property (5.3) is important when using the test defined by (5.2) for constructing confidence regions for the range of $P_o$. This is done in the next section.

Although the test of $H_o$ defined by (5.2) was intuitively motivated, it does have the following important properties.

THEOREM 5.3

(i) (5.2) is a consistent test of $H_o$. That is, if $H_o$ is not true, then Prob[rejecting $H_o$] $\to$ 1.

(ii) (5.2) is invariant under post-multiplication of A by a non-singular matrix.

PROOF. See the appendix.

Theorem 5.3.ii is important since the hypothesis $H_o$ is invariant under post-multiplication of A by a non-singular matrix. Thus, the test given by (5.2) tests whether the space spanned by the columns of A is a subspace of the range of $P_o$.

REMARK 1. If the assumption $\lambda_{i-1} \neq \lambda_i$ or the assumption $\lambda_{i+m-1} \neq \lambda_{i+m}$ is false, then the asymptotic chi-square test given by (5.2) is not generally valid. If the assumptions on the eigenvalues are true, then by Lemma 4.1, the "sample" size n necessary to insure that the asymptotic chi-square test is a "good" approximation is in general inversely related to the quantity $\min(\lambda_{i-1} - \lambda_i, \lambda_{i+m-1} - \lambda_{i+m})$. In addition, if $\lambda_{i-1}$ is "close" to $\lambda_i$, one may not wish to study the eigenspace associated with $\lambda_{i-1}$ separately from the eigenspace associated with $\lambda_i$. So, in practice, before determining which eigenspaces are of interest, a study of the eigenvalues would be desirable.

REMARK 2. Let $v = \{\lambda_j, j \in I\}$, where I is some index set. Under the assumption $\lambda_j \neq \lambda_k$ for all $j \in I$ and $k \notin I$, consider the hypothesis

(5.4) $H_o$: the columns of A lie in the subspace generated by the eigenvectors of M associated with $\{\lambda_j, j \in I\}$,

where A is $(p \times r)$ with $\text{rank}(A) = r \leq m = \text{rank}(\Sigma_{\lambda \in v} P_\lambda)$. This hypothesis can be tested by using the test given by (5.2) provided w is replaced by v and $\hat{w}$ is replaced by $\hat{v} = \{\hat{\lambda}_j, j \in I\}$.

REMARK 3. Under the assumption $\lambda_{i-1} \neq \lambda_i$ and $\lambda_{i+m-1} \neq \lambda_{i+m}$, consider the hypothesis

(5.5) $H_o$: the eigenvectors of M associated with the roots $\lambda_i, \lambda_{i+1}, \ldots, \lambda_{i+m-1}$ lie in the subspace generated by the columns of A, where A is $(p \times r)$ with $\text{rank}(A) = r \geq m$. This hypothesis can be tested by using the following approach. Let B be $[p \times (p-r)]$ with $\text{rank}(B) = p-r$ and such that $A'B = 0$. The hypothesis (5.5) can then be rephrased as

(5.6) $H_o$: the columns of B lie in the subspace generated by the eigenvectors of M' associated with the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_{i-1}, \lambda_{i+m}, \ldots, \lambda_p$.

Note that if M is symmetric in the metric of $\Gamma$, then M' is symmetric in the metric of $\Gamma^{-1}$. It is easy to verify that if the conditions on M, $\Gamma$, $M_n$, $\Gamma_n$, $\Sigma$ and $\Sigma_n$ given by Assumptions 3.1 and 3.2 are satisfied, then the conditions are satisfied when M, $\Gamma$, $M_n$, $\Gamma_n$, $\Sigma$ and $\Sigma_n$ are replaced by $M'$, $\Gamma^{-1}$, $M_n'$, $\Gamma_n^{-1}$, $\text{var}(N') = I_{(p,p)} \Sigma I_{(p,p)}$ and $I_{(p,p)} \Sigma_n I_{(p,p)}$ respectively. So, by remark 2, the results of this section apply to testing the hypothesis (5.6).

Note that if $r = m$, then the hypothesis (1.1) and (5.5) are equivalent. For this case, the test given by (5.2) when applied to the hypothesis (1.1) is the same as the test for (5.5) suggested in this remark.

## 6. ASYMPTOTIC CONFIDENCE REGIONS

For A which is $(p \times m)$, let L(A) represent the space spanned by the columns of A. That is,

$$L(A) = \{\underset{\sim}{v} \in R^p \mid \underset{\sim}{v} = A\underset{\sim}{w} \text{ for some } \underset{\sim}{w} \in R^m\}.$$

The test of hypothesis (1.1), when $r = m$, given by (5.2) yields the following asymptotic $(1-\alpha)$ confidence region for the range of $P_o$,

$$\{L(A) \mid A \text{ is } (p \times m),\ \text{rank}(A) = m, \text{ and } T_n(A) < \chi^2_{m(p-m)}(\alpha)\}. \tag{6.1}$$

One "undesirable" aspect of this confidence region is that $T_n(A)$ involves a generalized inverse of $\hat{\Sigma}_o(A)$, which must be recalculated for each A. However, this problem can be alleviated and the confidence region can be given a simplier representation.

To make the simplification, let

$$(6.2) \qquad X_n = [\hat{\underset{\sim}{x}}_i \; \hat{\underset{\sim}{x}}_{i+1} \; \cdots \; \hat{\underset{\sim}{x}}_{i+m-1}],$$

where $\{\hat{\underset{\sim}{x}}_j\}$ is defined such that $M_n\hat{\underset{\sim}{x}}_j = \hat{\lambda}_j\hat{\underset{\sim}{x}}_j$, and $\hat{\underset{\sim}{x}}_j'\Gamma_n\hat{\underset{\sim}{x}}_k = \delta_{ij}$. By noting that $\hat{P}_o = X_nX_n'\Gamma_n$, it can then be easily verified that

$$\Sigma_o(A) = (A'\Gamma_nX_n \otimes I)\Sigma_o(X_n)(X_n'\Gamma_nA \otimes I).$$

So, by Theorem 5.3.ii, if $\text{rank}(\hat{P}_oA) = m$, then

$$T_n(A) = T_n[A(X_n'\Gamma_nA)^{-1}] = a_n^2\{\text{vec}[A(X_n'\Gamma_nA)^{-1} - X_n]\}'\Sigma_o(X_n)^+\text{vec}[A(X_n'\Gamma_nA)^{-1} - X_n].$$

Thus, (6.1) can be rewritten as

$$(6.3) \quad \{L(A) \mid X_n'\Gamma_nA = I \text{ and } a_n^2[\text{vec}(A-X_n)]'\Sigma_o(X_n)^+\text{vec}(A-X_n) < \chi^2_{m(p-m)}(\alpha)\}.$$

For the special case m = 1, (6.3) reduces to

$$(6.4) \qquad \{c\underset{\sim}{a} \mid \hat{\underset{\sim}{x}}_i'\Gamma_n\underset{\sim}{a} = 1, \text{ and } a_n^2(\underset{\sim}{a}-\hat{\underset{\sim}{x}}_i)'\Lambda_n^+(\underset{\sim}{a}-\hat{\underset{\sim}{x}}_i) < \chi^2_{p-1}(\alpha)\},$$

$$\text{where } \Lambda_n = [\hat{\underset{\sim}{x}}_i' \otimes (M_n-\hat{\lambda}_iI)^+]\Sigma_n[\hat{\underset{\sim}{x}}_i \otimes (M_n'-\hat{\lambda}_iI)^+].$$

If $M_n$ and M are symmetric, (6.3) and (6.4) respectively reduce to

$$(6.5) \quad \{L(A) \mid X_n'A = I \text{ and } a_n^2[\text{vec}(A)]'\Sigma_o(X_n)^+[\text{vec}(A)] < \chi^2_{m(p-m)}(\alpha)\},$$

and

$$(6.6) \qquad \{c\underset{\sim}{a} \mid \hat{\underset{\sim}{x}}_i'\underset{\sim}{a} = 1, \text{ and } a_n^2\underset{\sim}{a}'\Lambda_n^+\underset{\sim}{a} < \chi^2_{p-1}(\alpha)\},$$

$$\text{where } \Lambda_n = [\hat{\underset{\sim}{x}}_i' \otimes (M_n-\hat{\lambda}_iI)^+]\Sigma_n[\hat{\underset{\sim}{x}}_i \otimes (M_n-\hat{\lambda}_iI)^+].$$

## 7. EXAMPLES

I) PRINCICAL COMPONENTS ANALYSIS. One of the most common uses of eigenvectors in statistics is in the principal components analysis of a covariance matrix for a multivariate normal distribution. For this case, $M_n$ is taken to be the sample covariance matrix from a sample of size n from a multivariate normal distribution with nonsingular covariance matrix M. That is, $M_n = (1/n)\Sigma_{i=1}^{n}(\underset{\sim}{Y}_i-\bar{\underset{\sim}{Y}})(\underset{\sim}{Y} - \bar{\underset{\sim}{Y}})'$, where $\underset{\sim}{Y}_1, \underset{\sim}{Y}_2, \ldots, \underset{\sim}{Y}_n$ are i.i.d. Normal$(\underset{\sim}{\mu}, M)$.

It is well-known that $\sqrt{n}(M_n-M) \to N$ in distribution, where vec(N) has a multivariate normal distribution with zero mean and covariance matrix $\Sigma = (I + I_{(p,p)})(M \otimes M)$. (See Izenman (1976) or Magnus and Neudecker (1979).) Choose $\Sigma_n = (I + I_{(p,p)})(M_n \otimes M_n)$. It can then be verified that Assumptions 3.1 and 3.2 are satisfied, and so the results of this paper apply to this example.

On the set rank$(\hat{P}_o A) = r$, we have the representation

$$T_n(A) = n \; \Sigma_{\mu \in \hat{w}} \; \mu^{-1}\mathrm{Trace}\{A'\hat{P}_\mu A[A'X_n D_n(\mu)X_n'A]^{-1}\}, \tag{7.1}$$

where $X_n$ is defined in (6.2) and $D_n(\mu)$ is an $(m \times m)$ diagonal matrix with entries $\hat{\lambda}_j/(\hat{\lambda}_j-\mu)^2$, $j = i, i+1, \ldots, i+m-1$. For the important case $r = m$, (7.1) becomes

$$T_n(A) = n \; \mathrm{Trace}\{A'M_n^{-1}A(X_n'A)^{-1}\Delta_n(A'X_n)^{-1} + A'MA(X'A)^{-1}\Delta_n^{-1}(A'X_n)^{-1} - 2A'A(X_n'A)(A'X_n)^{-1}\}, \tag{7.2}$$

where $\Delta_n$ is an $(m \times m)$ diagonal matrix with entries $\hat{\lambda}_i, \hat{\lambda}_{i+1}, \ldots, \hat{\lambda}_{i+m-1}$.

In particular, for $m = 1$, (7.2) becomes

$$T_n(\underset{\sim}{a}) = n[\hat{\lambda}_i \underset{\sim}{a}'M_n^{-1}\underset{\sim}{a} + \hat{\lambda}_i^{-1}\underset{\sim}{a}'M_n\underset{\sim}{a} - 2\underset{\sim}{a}'\underset{\sim}{a}]/(\underset{\sim}{a}'\hat{\underset{\sim}{x}}_i)^2 \tag{7.3}$$

Under the null hypothesis, this statistic is asymptotically equivalent to the test statistic given by Anderson (1963). Anderson's statistic is

$$n[\hat{\lambda}_i \underset{\sim}{a}' M_n^{-1} \underset{\sim}{a} + \hat{\lambda}_i \underset{\sim}{a}' M_n \underset{\sim}{a} - 2], \tag{7.4}$$

with $\underset{\sim}{a}$ normalized such that $\underset{\sim}{a}'\underset{\sim}{a} = 1$.

II) CANONICAL ANALYSIS. Another common use of eigenvectors in statistics is in the canonical analysis of the joint covariance matrix for two random vectors which are jointly multivariate normal. For this case, $M_n = \hat{C}_{11}^{-1}\hat{C}_{12}\hat{C}_{22}^{-1}\hat{C}_{21}$ and $M = C_{11}^{-1}C_{12}C_{22}^{-1}C_{21}$, where

$$C_n = \begin{bmatrix} \hat{C}_{11} & \hat{C}_{12} \\ \hat{C}_{21} & \hat{C}_{22} \end{bmatrix} \text{ and } C = \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix}.$$

$C_n$ represents the sample covariance matrix for a sample of size n from a $(p \times q)$ multivariate normal vector with nonsingular covariance matrix C. For this example, $\Gamma = C_{11}$ and $\Gamma_n = \hat{C}_{11}$.

By expanding $M_n$ in a Taylor series about C, we note that $\sqrt{n}(M_n - M) \to N$ in distribution, where vec(N) has a multivariate normal distribution with mean zero and covariance matrix

$$\begin{aligned} \Sigma = {} & \Gamma(I-M) \otimes \Gamma^{-1}M' + \Gamma(M-M^2) \otimes \Gamma^{-1}(I-2M') \\ & + I_{(p,p)}\{(I-M) \otimes (M'-[M']^2) + (M-M^2) \otimes (I-M')\} \end{aligned} \tag{7.5}$$

Choose $\Sigma_n$ to have the same form as $\Sigma$ with $\Gamma_n$ and $M_n$ replacing $\Gamma$ and M respectively. For $C_{12} \neq 0$, it can be verified that Assumptions 3.1 and 3.2 are satisfied, and so the results of this paper apply to this example.

On the set $\text{rank}(\hat{P}_o A) = r$, we have the representation

$$T_n(A) = n\, \Sigma_{\mu \notin \hat{w}} \text{Trace}\{A'\Gamma_n \hat{P}_\mu A[A'\Gamma_n X_n D_n(\mu) X_n' \Gamma_n A]^{-1}\} \tag{7.6}$$

where $D_n(\mu)$ is an $(m \times m)$ diagonal matrix with entries

$$(1-\hat{\lambda}_j)(\mu+\hat{\lambda}_j-2\mu\hat{\lambda}_j)/(\hat{\lambda}_j-\mu)^2, \quad j = i,i+1,\ldots,i+m-1.$$

In particular, for $m = 1$, (7.6) becomes

$$(7.7) \quad T_n(\underset{\sim}{a}) = n\underset{\sim}{a}'\Gamma_n(M_n-\hat{\lambda}_i I)^2[(1-2\hat{\lambda}_i)M_n + \hat{\lambda}_i I]^{-1}\underset{\sim}{a}/[(1-\hat{\lambda}_i)(\underset{\sim}{a}'\Gamma_n\hat{\underset{\sim}{x}}_i)^2].$$

APPENDIX

In this appendix, the proofs for section 5 are given. These proofs are given in the same order in which they appear in section 5 with the exception that the proof for Theorem 5.1 follows the proof for Theorem 5.2.

Before presenting the proofs, the following lemmas concerning quadratic forms involving singular matrices are needed.

LEMMA A.1. Let S be a positive semi-definite symmetric matrix of order $(s \times s)$, $\underset{\sim}{x} \in \text{range}(S)$, and B a $(s \times k)$ matrix with $\text{rank}(B) = k$, then (i) $\underset{\sim}{x}'B(B'SB)^{-}B'\underset{\sim}{x}$ is invariant with respect to the choice of the generalized inverse for $B'SB$.

(ii) $\underset{\sim}{x}'B(B'SB)^{-}B'\underset{\sim}{x} \le \underset{\sim}{x}'S^{-}\underset{\sim}{x}$, with equality if $k = s$.

(iii) $\underset{\sim}{x}'S^{-}\underset{\sim}{x} \ge (\underset{\sim}{x}'\underset{\sim}{x})^{2}(\underset{\sim}{x}'S\underset{\sim}{x})^{-1}$.

LEMMA A.2. For $\underset{\sim}{Y}_n$, an s-dimensional random vector, if

(i) $\underset{\sim}{Y}_n \to \text{Normal}(0,S)$ in distribution, with $\text{rank}(S) = r$,

(ii) $S_n \to S$ in probability,

(iii) $\text{rank}(S_n) \to r$ in probability, and

(iv) $\text{Prob}[\underset{\sim}{Y}_n \in \text{range}(S_n)] \to 1$,

then for any sequence of generalized inverses for $S_n$,

$\underset{\sim}{Y}_n'S_n^{-}\underset{\sim}{Y}_n \to \chi_r^2$, in distribution.

LEMMA A.3. Let $S_n$ and S be random matrices such that $S_n \to S$ in distribution. If $\text{rank}(S) = r$ almost surely, and $\text{rank}(S_n) \le r$, then $\text{rank}(S_n) \to r$ in probability. Alternatively, $\text{Prob}[\text{rank}(S_n) = r] \to 1$.

The proof for Lemma A.1 is straight forward and can be found in Appendix C of the author's dissertation. Lemma A.2 is a corrected version of Theorem 1.b given by Moore (1977). Proof's for Lemmas A.2 and A.3 can also be found in Appendix C of the author's dissertation.

Finally, the following special case of the theorem given by Okamoto (1973) is needed before the proofs for section 5 can be given.

LEMMA A.4. If B is a $(k_1 \times k_2)$ random matrix such that $\mathrm{vec}(B) \sim \mathrm{Normal}(\underset{\sim}{0}, S)$ with $\mathrm{rank}(S) = k_1 k_2$, then $\mathrm{rank}(B) = \min(k_1, k_2)$ almost surely.

<u>Proof of Theorem 5.2.i.</u> By noting that $\hat{\Sigma}_o(A) = \hat{\Sigma}_o(\hat{P}_o A)$ is a sample version of $\Sigma_o(A)$ in Theorem 4.2, we obtain

(A.1) $$\mathrm{rank}[\hat{\Sigma}_o(A)] = (p-m)r,$$

whenever $\mathrm{rank}(\hat{P}_o A) = r$. It is then easy to verify that if $\mathrm{rank}(\hat{P}_o A) = r$, then $\mathrm{range}[\hat{\Sigma}_o(A)] = \mathrm{range}[I \otimes (I-\hat{P}_o)]$. So, by noting that $[I \otimes (I-\hat{P}_o)]\mathrm{vec}[(I-\hat{P}_o)A] = \mathrm{vec}[(I-\hat{P}_o)A]$, we have

(A.2) if $\mathrm{rank}(\hat{P}_o A) = r$, then $\mathrm{vec}[(I-\hat{P}_o)A]$ is in $\mathrm{range}[\hat{\Sigma}_o(A)]$.

The proof for Theorem 5.2.i is completed by applying Lemma A.1.i to $T_n(A)$, using $B = I$.

<u>Proof of Theorem 5.2.ii.</u> Let $r_1 = \mathrm{rank}(P_o A)$, and let $G_1$ and $\hat{G}_1$ be the total eigenprojections associated with the $r_1$ largest eigenvalues of $A'P_o'\Gamma P_o A$ and $A'\hat{P}_o'\Gamma\hat{P}_o A$ respectively. Also, let $G_o = I-G_1$, and $\hat{G}_o = I-\hat{G}_1$.

By noting that $(\hat{P}_o A\hat{G}_1)'\Gamma(\hat{P}_o A\hat{G}_o) = 0$ and $(\hat{P}_o A\hat{G}_o)(\hat{P}_o A\hat{G}_1)' = 0$, we have $\mathrm{rank}(\hat{P}_o A) = \mathrm{rank}(\hat{P}_o A\hat{G}_o) + \mathrm{rank}(\hat{P}_o A\hat{G}_1)$. Now, since $\mathrm{rank}(\hat{P}_o A\hat{G}_1) = \mathrm{rank}(\hat{G}_1' A'\hat{P}_o'\Gamma\hat{P}_o A\hat{G}_1)$ and $\hat{G}_1' A'\hat{P}_o'\Gamma\hat{P}_o A\hat{G}_1 \to A'P_o'\Gamma P_o A$ in probability, it follows from the continuity of eigenvalues of symmetric matrices that

$\text{rank}(\hat{P}_o A\hat{G}_1) \to r_1$ in probability. So, we only need to show that

$\text{rank}(\hat{P}_o A\hat{G}_o) \to (r-r_1)$ in probability.

To show this, using Lemma 4.1, we obtain the Taylor series

$$\hat{G}_o = G_o - G_o(A'\hat{P}_o'\Gamma\hat{P}_o A - A'P_o'\Gamma P_o A)(A'P_o'\Gamma P_o A)^+$$
$$- (A'P_o'\Gamma P_o A)^+(A'\hat{P}_o'\Gamma\hat{P}_o A - A'P_o'\Gamma P_o A)G_o + O(a_n^{-2}).$$

Using this expansion, and the expansion for $\hat{P}_o$ given by Lemma 4.1, we obtain

(A.3) $$a_n(\hat{P}_o A\hat{G}_o) \to W = -B_o\Sigma_{\lambda\epsilon w}P_\lambda NQ_\lambda AG_o$$

in distribution, where $B_o = [P_o - P_oA(A'P_o'\Gamma P_oA)^+A'P_o'\Gamma]$ .

Noting that $G_o$ and $B_o$ are projections with $\text{rank}(G_o) = (r-r_1)$ and $\text{rank}(B_o) = (m-r_1)$, we can choose matrices $C_1$ and $C_2$ of dimension $[r \times (r-r_1)]$ and $[p \times (m-r_1)]$ respectively, such that $\text{rank}(C_1) = (r-r_1)$, $\text{rank}(C_2) = (m-r_1)$, $G_oC_1 = C_1$ and $B_oC_2 = C_2$. It can then be shown that for any $[(r-r_1) \times (m-r_1)]$ matrix C,

(A.4) $$\text{var}\{[\text{vec}(C)]'\text{vec}(C_2'\Gamma WC_1)\} = [\text{vec}(\Gamma C_o)]'\Sigma\text{vec}(\Gamma C_o),$$

where $C_o = \Sigma_{\lambda\epsilon w}P_\lambda C_2CC_1'AQ_\lambda'$. For $C \neq 0$, (A.4) is positive. This follows by noting that $\Sigma_{\lambda\epsilon w}\Sigma_{\mu\notin w}\mu^{-1}(\mu-\lambda)P_\mu M(C_o+C_o')P_\lambda' = AC_1CC_2' \neq 0$, and so $M(C_o+C_o') \neq 0$. Thus, by Assumption 3.1.iii, (A.4) is not zero.

Since (A.4) is positive for $C \neq 0$, this implies that $\text{rank}[\text{var}(C_2'\Gamma WC_1)] = (m-r_1)(r-r_1)$. So, by Lemma A.4, $\text{rank}(C_2'\Gamma WC_1) = (r-r_1)$ almost surely. Also, since $(r-r_1) \geq \text{rank}(W) \geq \text{rank}(C_2'\Gamma WC_1)$, the rank of W is almost surely equal to $(r-r_1)$.

The proof to Theorem 5.2.ii is completed by noting that $\mathrm{rank}(a_n \hat{P}_o A \hat{G}_o) \le (r-r_1)$, and then applying Lemma A.3.

Proof of Theorem 5.1. This proof consists of showing that the conditions of Lemma A.2 are satisfied when using $a_n \mathrm{vec}[(\hat{P}_o - I)A]$ for $\underset{\sim}{Y}_n$ and $\hat{\Sigma}_o(A)$ for $S_n$. Condition (i) follows from (A.2). Condition (iv) follows from (A.2) and Theorem 5.2.ii. If condition (ii) is satisfied, then condition (iii) follows from (A.1) and Lemma A.3. So, it only needs to be shown that condition (ii) is satisfied, that is, to show that $\hat{\Sigma}_o(A) \to \Sigma_o(A)$ in probability. Since $\Sigma_n \to \Sigma$ in probability, it is sufficient to prove that

$$\text{(A.5)} \qquad \Sigma_{\lambda \in \hat{w}} \hat{P}_\lambda' \otimes \hat{Q}_\lambda (I - \hat{P}_o) \to \Sigma_{\lambda \in w} P_\lambda' \otimes Q_\lambda (I - P_o),$$

in probability.

Before probing (A.5), additional notation is needed. If $\lambda_{a-1} \ne \lambda_a = \ldots = \lambda_{a+b} \ne \lambda_{a+b+1}$, then define $l(\lambda_a) = a$, $u(\lambda_a) = a+b$, $m(\lambda_a) = b+1$ and $\hat{w}(\lambda_a) = \{\hat{\lambda}_a, \ldots, \hat{\lambda}_{a+b}\}$. Also, define $d(\lambda,\mu) = (\lambda-\mu)^{-1}$, and for $\lambda$ and $\mu$ which are eigenvalues of M, define

$\hat{d}(\lambda,\mu) = [\Sigma_{j=l(\lambda)}^{u(\lambda)} \Sigma_{k=l(\mu)}^{u(\mu)} d(\hat{\lambda}_j, \hat{\lambda}_k)]/[m(\lambda)m(\mu)]$, provided $\lambda \ne \mu$. By the continuity of eigenvalues, that is Lemma 2.1, we note that $\hat{d}(\lambda,\mu) \to d(\lambda,\mu)$ in probability. Also, if $\alpha \in \hat{w}(\lambda)$ and $\beta \in \hat{w}(\mu)$, then $d(\alpha,\beta) \to d(\lambda,\mu)$ in probability.

So, if we define the norm on all $(p \times p)$ matrices $||B||_G = [\max \text{ eigenvalue of } G^{-1}B'GB]^{\frac{1}{2}}$, where G is a symmetric positive definite matrix of order $(p \times p)$, then

$$\text{(A.6)} \quad ||\Sigma_{\alpha \in \hat{w}(\lambda)} \Sigma_{\beta \in \hat{w}(\mu)} d(\alpha,\beta) \hat{P}_\alpha' \otimes \hat{P}_\beta - \hat{d}(\lambda,\mu)[\Sigma_{\alpha \in \hat{w}(\lambda)} \hat{P}_\alpha' \otimes \Sigma_{\beta \in \hat{w}(\mu)} \hat{P}_\beta]||\Gamma_n$$

$$\le \Sigma_{j=l(\lambda)}^{u(\lambda)} \Sigma_{k=\ (\mu)}^{u(\mu)} |d(\hat{\lambda}_i, \hat{\lambda}_k) - \hat{d}(\lambda,\mu)|.$$

By the arguments of the previous paragraph, it follows that the right-hand side of (A.6) goes to zero in probability. Note that if $||B_n||_{\Gamma_n} \to 0$ in probability, then $||B_n||_{\Gamma} \to 0$ in probability, since $\Gamma_n \to \Gamma$ in probability. So, the left-hand side of (A.6) goes to zero in probability if the norm $||\cdot||_{\Gamma_n}$ are replaced by $||\cdot||_{\Gamma}$.

Thus, since by (4.2) $\Sigma_{\alpha\epsilon\hat{w}(\lambda)}\hat{P}_{\alpha} \to P_{\lambda}$ in probability, we have

$$\text{(A.7)} \quad \Sigma_{\lambda\epsilon w}\Sigma_{\mu\notin w}\Sigma_{\alpha\epsilon\hat{w}(\lambda)}\Sigma_{\beta\epsilon\hat{w}(\mu)}d(\alpha,\beta)\hat{P}'_{\alpha} \otimes \hat{P}_{\beta} \to \Sigma_{\lambda\epsilon w}\Sigma_{\mu\notin w}d(\lambda,\mu)P'_{\lambda} \otimes P_{\mu} ,$$

in probability. The proof of Theorem 5.1 is completed by noting that the left and right hand sides of statement (A.7) are the same as the left and right hand sides of statement (A.5) respectively.

<u>Proof of statement (5.3).</u> Let $r_o = \mathrm{rank}(\hat{P}_oA)$, and let B be a $[r \times (r-r_o)]$ matrix with $\mathrm{rank}(B) = r-r_o$ and such that $\hat{P}_oAB = 0$. By (A.2), we can apply Lemma A.1.ii and A.1.iii to obtain

$$\text{(A.8)} \qquad T_n(A_k) \geq T_n(A_kB) \geq a_n^2(\underset{\sim}{b}'_k\underset{\sim}{b}_k)^2[\underset{\sim}{b}'_k\hat{\Sigma}_o(A_kB)\underset{\sim}{b}_k]^{-1} ,$$

where $\underset{\sim}{b}_k = \mathrm{vec}[(\hat{P}_o-I)A_kB]$. As $k \to \infty$, $\hat{\Sigma}_o(A_kB) \to 0$ and $\underset{\sim}{b}_k \to \mathrm{vec}(-AB)$, which is non-zero. Thus, the right-hand side of (A.8) goes to infinity.

<u>Proof of Theorem 5.3.i.</u> By Theorem 5.2.i, $\mathrm{Prob}[\mathrm{rank}(\hat{P}_oA) < r] \to 0$. Thus, we only need to show that if $H_o$ is false, then $\mathrm{Prob}[\mathrm{rank}(\hat{P}_oA) = r,\ T_n(A) > c] \to 1$, for any constant c. By (A.2), we can apply Lemma A.1.iii to obtain

$$\text{(A.9)} \qquad T_n(A) \geq a_n^2(\underset{\sim}{c}'_n\underset{\sim}{c}_n)^2[\underset{\sim}{c}'_n\hat{\Sigma}_o(A)\underset{\sim}{c}_n]^{-1} ,$$

where $\underset{\sim}{c}_n = \mathrm{vec}[(\hat{P}_o-I)A]$. Note that the proof for $\hat{\Sigma}_o(A)$ converging to $\Sigma_o(A)$ in probability given in the proof for Theorem 5.1 does not depend on the truth of $H_o$. So, since $\underset{\sim}{c}_n$ converges in probability to vec$[(P_o-I)A]$, which is non-zero, it follows that the probability the right-hand side of (A.9) is greater than any fixed constant c goes to one.

<u>Proof of Theorem 5.3ii.</u> Let B be an $(r \times r)$ non-singular matrix. If rank$(\hat{P}_o A) = r$, then rank$(\hat{P}_o AB) = r$. Also, by (A.2) and Lemma A.1.i, we have $T_n(AB) = T_n(A)$.

## REFERENCES


Anderson T. W. (1951). An asymptotic distribution of certain characteristic roots and vectors. Proc. Second Berkeley Symp. Math. Stat. Probability 103-130. Univ. California Press.

Anderson, T. W. (1963). Asymptotic theory for principal components analysis. Ann. Math. Stat. 34 122-148.

Chambers, J. M. (1967). On methods of asymptotic approximations for multivariate distributions. Biometrika 54 367-383.

Izenman, A. J. (1976). Reduced-rank regression for the multivariate linear model. J. Multivariate Analysis 5 248-264.

James, A. T. (1977). Test for a prescribed subspace of principal components. Multivariate Analysis - IV (P. R. Krishnaiah, Ed.) pp. 73-77. North Holland Pub. Co., Amsterdam.

Kato, T. (1966). Perturbation Theory for Linear Operators. Springer - Verlag, Berlin.

Mallows, C. L. (1961). Latent vectors of random symmetric matrices. Biometrika 48 133-149.

Magnus, J. R. and Neudecker, H. (1979). The commutation matrix: some properties and applications. Ann. Statist. 7 381-394.

Moore, D. S. (1977). Generalized inverses, Wald's method, and the construction of chi-squared tests of fit. J. Amer. Statist. Assoc. 72 131-137.

Nerring, E. D. (1970). Linear Algebra and Matrix Theory, 2nd Edition. Wiley, New York.

Neudecker, H. (1968). The Kronecker matrix product and some of its applications. Statistica Neerlandia 22 69-82.

Okamoto, M. (1973). Distinctness of the eigenvalues of a quadratic form in a multivariate normal sample. Ann. Statist. 1 763-765.

Rao, C. R. and Mitra, S. K. (1971) Generalized Inverses of Matrices and its Application. Wiley, New York.

Suguira, N. (1976). Asymptotic expansions of the distributions of the latent roots and latent vectors of the Wishart and multivariate F matrices. J. Multivariate Anal. 6 500-525.

Tyler, D. E. (1979). Redundancy Analysis and associated asymptotic distribution theory. Ph.D. dissertation, Princeton University, Princeton, New Jersey.

David E. Tyler
Department of Mathematical Sciences
Old Dominion University
Norfolk, Virginia 23508

## REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

| | | |
|---|---|---|
| 1. REPORT NUMBER<br>Tech. Report #157, Series 2 | 2. GOVT ACCESSION NO. | 3. RECIPIENT'S CATALOG NUMBER |
| 4. TITLE (and Subtitle)<br>ASYMPTOTIC INFERENCE FOR EIGENVECTORS | | 5. TYPE OF REPORT & PERIOD COVERED<br>Tech. Report |
| | | 6. PERFORMING ORG. REPORT NUMBER |
| 7. AUTHOR(s)<br>David E. Tyler | | 8. CONTRACT OR GRANT NUMBER(s)<br>N0014-67-0151-0017 and N00014-75-C-0453 |
| 9. PERFORMING ORGANIZATION NAME AND ADDRESS<br>Department of Statistics<br>Princeton University<br>Princeton, N.J. | | 10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS |
| 11. CONTROLLING OFFICE NAME AND ADDRESS<br>Office of Naval Research (Code 436)<br>Arlington, Virginia 22217 | | 12. REPORT DATE<br>October 1979 |
| | | 13. NUMBER OF PAGES<br>25 |
| 14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) | | 15. SECURITY CLASS. (of this report)<br>UNCLASSIFIED |
| | | 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE |

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

David Tyler is now with the Department of Mathematical Sciences, Old Dominion University, Norfolk, Virginia 23508

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

79 12 14 078

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

Asymptotic procedures are given for testing certain hypotheses concerning eigenvectors and for constructing confidence regions for eigenvectors. These asymptotic procedures are derived under fairly general conditions on the estimates of the matrix whose eigenvectors are of interest. Applications of the general results to principal components analysis and canonical variate analysis are given.